

Securing the Metaverse - Virtual Worlds Need REAL Governance

Kavya Pearlman
XR Safety Initiative
San Francisco, CA
kavya@xr.si.org

Sam Visner
The MITRE Corporation
McLean, Virginia
svisner@mitre.org

Marco Magnano
XR Safety Initiative
San Francisco Bay Area, CA
marco@xr.si.org

Ryan Cameron
XR Safety Initiative
San Francisco Bay Area, CA
ryancameron@xr.si.org

ABSTRACT

As the world becomes increasingly connected, and immersive technologies gain wider adoption both within government, businesses, and in the consumer market, a framework for security and privacy for these devices is required. From headsets to other wearables and related sensors, eXtended Reality (XR) technologies are now capable of gathering untold quantities of biometric data about users, potentially everything from a user's location and skin color to their eye and hand positions at any given time. As we build the next version of the Internet/Web 3.0, also known as "The Metaverse," it is critically important that we take a proactive approach and address some fundamental design choices about the principles of how we want it to operate, and potentially replicate or deepen what is broken about the Web today. Among the most important issues to address is data ownership and accountability. The National Institute of Standards and Technology (NIST) has offered basic guidance, while regional laws such as General Data Protection Regulations (GDPR), Children's Online Privacy Protection Rule (COPPA), and Family Educational Rights and Privacy Act (FERPA) govern some forms of data in specific locations. Despite the existing guidelines and regional laws, comprehensive protections are not in place to protect individuals and stakeholders in the Metaverse. With this in mind, the XR Safety Initiative (XRSI) proposes a Privacy and Safety Framework that sets a baseline set of standards, guidelines, and best practices that are regulation agnostic. It incorporates privacy requirements drawn from the GDPR, NIST guidance, FERPA, COPPA, and other evolving laws. The framework is designed to adapt and include novel requirements as new regulations come into effect. With Version 1.1 expected to be published in 2021, this paper provides an overview of the framework, how it was developed, and highlights changes in version 1.1, and most importantly, provides a baseline for securing the Metaverse. We also discuss who can benefit from this framework and offer guidance to organizations, developers, and service providers on how to implement the framework for added security and privacy designed into their product or service. This paper also provides organizations with consumer-facing Modeling and Simulation (M&S) use cases, an understanding of what a security posture should include – beyond traditional authorities to operate (ATO's) – as more organizations look to adopt emerging technologies such as XR and bring the Metaverse to life.

VISIT...

LANZAROTE
Caliente.COM

ABOUT THE AUTHORS

Kavya Pearlman Well known as the “Cyber Guardian,” founder and CEO of the XR Safety Initiative (XRSI), Kavya Pearlman is an award-winning cybersecurity professional with a deep interest in immersive and emerging technologies. She recently launched a novel XRSI Privacy and Safety Framework for the XR domain. She has been named one of the Top Cybersecurity influencers for three consecutive years, 2018-2019-2020 by IFSEC Global. Kavya has previously advised Facebook on third-party security risks during the 2016 US presidential elections and worked as the head of security for the oldest virtual world, “Second Life” by Linden Lab. Kavya is the co-host of the immersive podcast “Singularity Watch” and one of the Top 50 speakers in the cybersecurity industry. Kavya has founded The CyberXR Coalition that now focuses on diversity and inclusion and the cross-section of Cybersecurity and XR, helped launch a trustworthy XR news platform, ReadyHackerOne, and established a Medical XR Advisory Council.

Samuel Sanders Visner is a Tech Fellow at the MITRE Corporation and former Director of the National Cybersecurity Federally Funded Research and Development Center (MITRE), sponsored by the National Institute of Science and Technology (appointed October 2017). Sam is Director of Security, Netcracker Technology. Sam also serves as a member of the Cyber Council of the Intelligence and National Security Alliance and is a former member of the Cyber Committee of the Armed Forces Communications and Electronics Association. Sam is an adjunct professor of Science and Technology in International Affairs at Georgetown University, where he teaches a course on cybersecurity policy, operations, and technology. Sam is also a member of the Council on Foreign Relations and the Atlantic Council and is a member of the Intelligence Community Studies Board, sponsored by the National Academy of Science and serving the Office of the Director of National Intelligence. Sam is also a former member of the Army Science Board. Sam served previously as Senior Vice President at ICF (General Manager, Cybersecurity and Resilience), Vice President at CSC (General Manager, CSC Global Cybersecurity), Senior Vice President at SAIC, and as Chief of Signals Intelligence Programs at the National Security Agency, from which he received the Agency's highest award for civilian service. Sam also served as a member of the Board of Directors, CVG/Avtec (2008-2010). Sam served twice on the Intelligence, Surveillance, and Reconnaissance Task Force of the Defense Science Board and has published articles on national and cybersecurity in World Politics Review, the Georgetown Journal of International Affairs, and the Defense Intelligence Journal. Sam is married to Antoinette (Toni) Burnham, Washington DC's leading urban beekeeper.

Marco Magnano, Executive Director of Communication at XRSI, is a multimedia journalist with ten years of live radio experience and tremendous working knowledge of advanced media technologies. Marco is based in Europe and a solid strategic and exclusive media contributor for XRSI, building a dedicated awareness space via the ReadyHackerOne platform. He assembled his first Personal Computer (PC) at the age of Four. Thirty years later, the mindset of building new ideas and deconstructing old ones is still there in his DNA.

XRSI's Executive Director of Communications, Marco Magnano, speaks English, French, Spanish, Portuguese, Italian, a bit of German and most recently learning Arabic. In recent years Marco has traveled across continents to uncover stories of various dimensions and domains, from the Syrian war crisis, challenges of migrations in Asia, Africa, and Europe to advocating for freedom of information and protecting the most vulnerable people. Marco is also the co-founder and creator of XRSI's trustworthy journalism platform, Ready Hacker One (RH1).

Ryan Cameron a founding contributor to the Second Life Grid Architecture working group and platform architect of multiple award winning enterprise-level software for early literacy and human skill acquisition and adoption at fortune 100 companies since 1997. Ryan is an adviser to the XR Safety Initiative (XRSI) and the co-chair of the Medical XR privacy and safety framework by XRSI. Ryan is currently developing a patent-pending standards platform for XR research and assessment to have a common architecture and user interface across multiple medical disciplines while using XR to engage patients in their health and humanity.

Introduction

The metaverse represents a new paradigm, perhaps the next phase in human evolution. Our bodies are projected in the form of avatars into a world we perceive similarly to how we perceive our own “real” world. The very concept of reality is challenged by the term “eXtended Reality”. It follows, then, that hard-fought rights we have achieved in our flesh and bone society will be as or even more critical in the Metaverse. We make this assertion because the very “digital air” we breathe, the “ground” we walk on, how we appear to others, what we perceive of others, what we feel, and how we are felt, can be directly and atomically controlled by the host and its army of algorithms. The Metaverse has typically been considered in the context of AR (Augmented Reality), VR (Virtual Reality) and Mixed Reality (MR), known collectively as Extended Reality (XR) becoming ubiquitous. Technically, Extended Reality (XR) is a fusion of all the simulated realities—including Augmented Reality (AR), Virtual Reality (VR), and Mixed Reality (MR)—which consists of technology-mediated experiences enabled via a broad spectrum of hardware and software, including sensory interfaces, applications, and infrastructures. XR is often referred to as immersive video content, enhanced media experiences, and interactive and multi-dimensional human experiences. With the mass adoption of XR and the Metaverse in the making, technologies such as Brain-computer interfaces (BCI) and autonomous systems enabled via Machine Learning increasingly interact with XR devices to produce astonishing virtual worlds. These mirror worlds built by extending realities and fueled by massive amounts of data are going to shape many different industries and missions, including many highly sensitive and regulated ones, such as healthcare and the military. Due to the uncharted and evolving nature of these brave new worlds, individuals and organizations are currently not fully aware of the irreversible and unintended consequences of XR on the digital and physical world (O’Brocháin et al., 2016). With the mass adoption of emerging technologies, including the US Army’s procurement of the HoloLens 2-based IVAS program (*U.S. Army to Use HoloLens Technology in High-Tech Headsets for Soldiers*, 2021), it is imperative to understand privacy and safety concerns and address them proactively. This framework by the XR Safety Initiative (XRSI) provides a baseline approach enabling better engineering practices that support privacy-by-design concepts and help organizations protect data and safeguard the XR platforms, applications, associated ecosystems and the Metaverse itself. The framework is the work of several interdisciplinary experts and serves as a tool for improving privacy, security, and safety through the human-centric design, pragmatic decision making, and proactive risk management.

According to XRSI’s Metaverse Safety Adviser, Doug Thompson, there are competing visions for the Metaverse and regardless of the direction in which it evolves, both centralized and decentralized architectures will co-exist in competition. The Metaverse is generally defined as a persistent, shared, 3-dimensional space. This definition arose primarily from popular culture[20][21]. The Metaverse has been called the ‘successor state’ to mobile computing. Just as mobile phones didn’t replace desktop computers, the Metaverse will increasingly come to dominate user time and attention. Instead of scrolling through mostly 2-dimensional screens, users will spend more and more of their time in digital experiences that are 3-dimensional, immersive, social, and increasingly indistinguishable from ‘reality’.

What is a Metaverse

“In practice, however, the Metaverse will evolve based on the convergence of several enabling technologies and experiences: AR, VR, 5G, AI, edge networks, improved graphics and computer hardware, improved mobile device capabilities. As these technologies and experiences reach more and more people, there is significant investment and grassroots pressure to make these technologies interoperable. Instead of downloading one client for Roblox and one for Minecraft, a user will be able to move with relative ease between these worlds”, adds Doug Thompson.

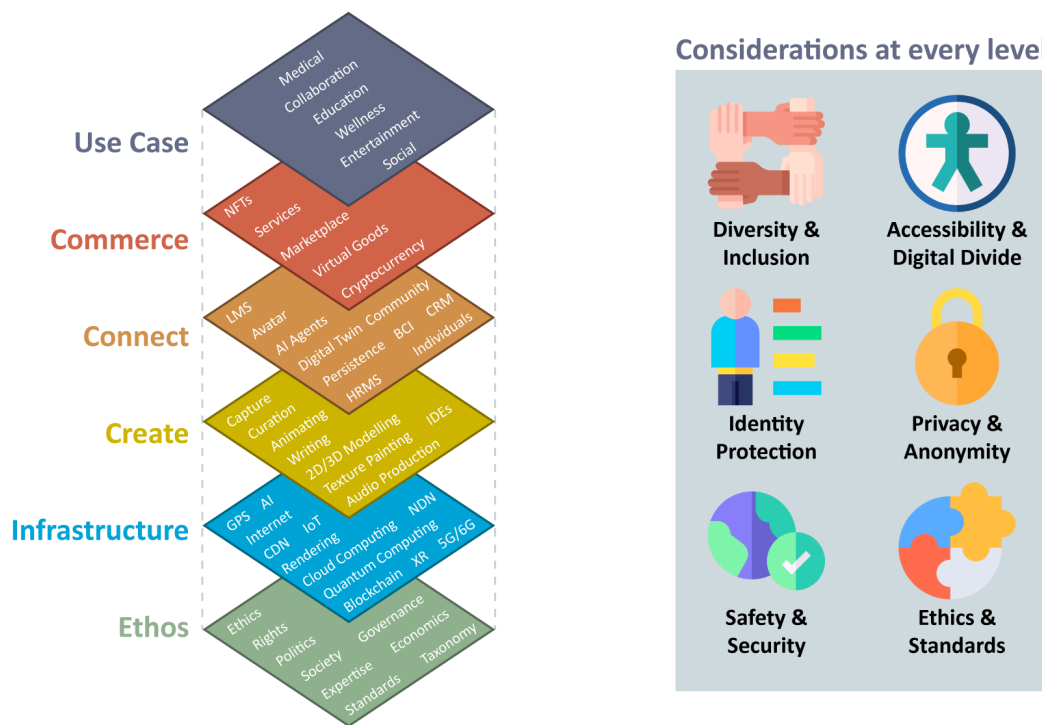


Figure 1 The Anatomy of the Metaverse

Therefore, the Metaverse represents interoperability between diverse spatial systems allowing for the flow of individuals, data, commerce, and culture across 3D spaces which may be owned by different actors. Clearly, the Metaverse is the new internet, where humans will do a lot more than just gaming, as described below:

Impact of the Metaverse on Humanity

Due to the catalyzing event of COVID-19, XR is already affecting every aspect of our lives and we are about to see exponential growth due to big commitments and promises being made by technology companies and overall convergence of various technologies such as BCI, AI/ML, 5G/6G, Robotics, Edge Computing, Decentralized Ledger Technologies (DLT) etc.

One of the key approaches to understanding the impact of the Metaverse on humanity is to look at the anatomy of any given instance of the Metaverse as a platform at its various levels of the stack as shown in the figure:

The metaverse will allow humans to create, connect, conduct commerce and even get medical assessment, diagnosis, treatment, and therapy by using various converging technologies. This creates significant challenges and raises questions that need to be addressed.

- **Diversity and Inclusion** - If we are going to extend reality via immersive technologies to live and work within the Metaverse, it is essential that the new realities are representative and considerate of who we really are as humans.
- **Accessibility and Digital Divide** - The Metaverse, in most cases will be accessible by reasonably priced Head Mounted Displays (HMDs) and XR devices, that can not only present accessibility challenges but also increase the digital divide.

- Identity Protection - As more individuals and organizations adopt the Metaverse for various use cases, we will need protection measures for minorities so that people are not targeted , harassed , or murdered for who they are , what they look like, what they believe in or how they pray. Clearly, new rules need to be created that go beyond some mandated login requirements in the name of safety consideration.
- Privacy and Anonymity - Given there can be no anonymity in the Metaverse, awareness is needed so people don't put themselves at risk in the immersive environments. Privacy itself needs redefining and clear understanding all the way to regulatory levels.
- Safety and Security - Current security frameworks and models do not take in account the expanded attack surface that goes beyond just computer nodes, servers, and the network. Therefore, safety considerations that account for human brain, human psychology, surrounding physical environments, etc. need to be factored into the new models for safety and security in the Metaverse.
- Ethics and Standards - To enable persistence, interoperability for the metaverse, new standards need to be established. These standards need to go beyond basic etiquettes online and have clear and specific ethical guidelines for various use cases within the metaverse, making “ethics” as the foundation of the Metaverse.

How We Got Here

From headsets to other wearables and related sensors, XR technologies are now capable of gathering enormous quantities of user biometric data either by direct data stream or by inferring from combined data sources, potentially including everything from a person's location and skin color to their eye and hand positions at any given time. Some, if not most, of this information has tremendous value to individuals and to organizations like the US military, but comprehensive regulations are not in place to protect the Metaverse stakeholders including consumers, private and public organizations, researchers, and governments etc. The National Institute of Standards and Technology (NIST) has offered basic guidance, while regional laws such as General Data Protection Regulation (GDPR), Children's Online Privacy Protection Act (COPPA), and the Family Educational Rights and Privacy Act (FERPA) govern some forms of data in specific contexts. XRSI's framework integrates them, adopting a more comprehensive approach that ties together regulations, practices, threat analyses, and specific use cases. The framework goes beyond the regulations and provides a greater understanding of these complex and emerging ecosystems. The XRSI Framework is not intended to be used as a compliance checklist, but as a baseline risk assessment model for building trust in the Metaverse. This approach strongly relies on a deep understanding of the least visible layer of the immersive experiences: data collection and usage.

Data Collection in Immersive Technologies

Many organizations are developing immersive technologies to build all-day wearable glasses that are spatially aware with the goal of delivering AR and VR experiences more immersive and integrated into the physical world serving as the baseline for the Metaverse. In order to achieve this outcome via the use of AI algorithms, a large amount of data collection is necessary. This is where most organizations and regulators face what we consider a Collingridge dilemma.

*“The social consequences of a technology cannot be predicted early in the life of the technology.
When change is easy, the need for it cannot be foreseen; when the need for change is apparent,
change has become expensive, difficult and time-consuming.”*

-David Collingridge

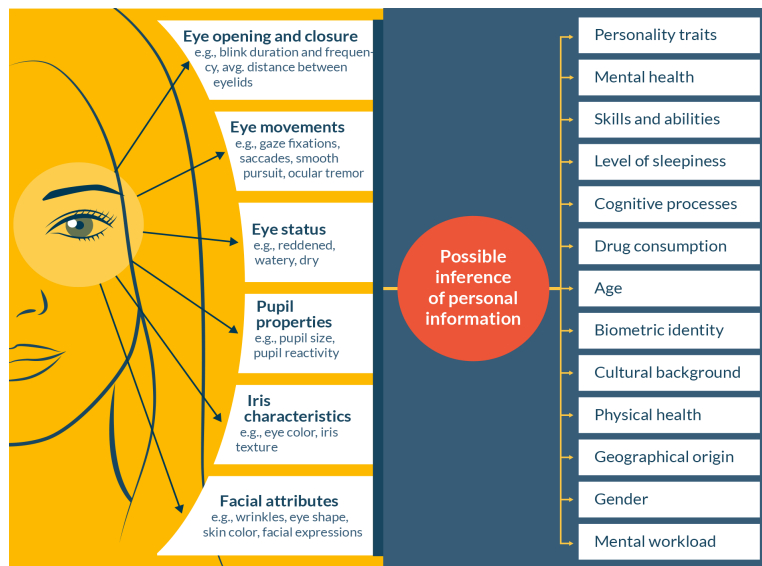


Figure 2 Possible uses of data collected by XR technologies. Credit: XRSI Framework.

According to the Collingridge dilemma, just like any other emerging technology research and development – the Metaverse faces a paradox or double-bind problem:

- An information problem: impacts cannot be easily predicted until the technology is extensively developed and widely used.
- A power problem: control or change is difficult when the technology has become fully entrenched. (Tech Liberation, 2021).

The dilemma equally applies to data collection in the Metaverse. Data is the heartbeat of the Metaverse. Meaningful experiences can only be realized when we use data *about* an individual *for* that individual's experience. This is the immersive nature of the technology. This inherently means developers need to know information about people – where you are, what you're looking at, if you're moving or not, etc. Developers may likely opt for a “more is better” approach when it comes to data collection. But the concerns for excessive data collection are heightened because of the large amount of real-time data collection and the potential inferences that are possibly made. While some inferences are necessary and even welcomed by individuals such as curated and customized shopping preferences many others such as Biometrically-Inferred Data (BID) are not and may cause harm to humans. BID is a collection of datasets resulting from information inferred from behavioral, physical, and psychological biometric identification techniques, and other nonverbal communication methods. For example, XR devices can lead to inferences such as biometric and gender identity, mental workload, mental health status, cognitive abilities, religious and cultural background, physical health, geographic origin, and many other skills, abilities, personality traits, and more. Based on different jurisdictions, organizations may be mandated to protect BID and require adopting a data governance framework like XRSI's. This will prevent excessive and unwarranted data collection at the hardware, operating system, API, and software level, leading to responsible research and innovation in the XR domain. Below are a few examples of the data tracked and collected by and for XR devices.

Data Collection Examples

A 2018 research study by Stanford University highlighted that “with VR, in addition to recording personal data regarding people's location, social ties, verbal communication, search queries, and product preferences, technology companies will also collect nonverbal behavior—for example, users' posture, eye gaze, gestures, facial expressions, and interpersonal distance” (Bailenson, 2018). According to the eminent psychologist Paul Ekman, “actions speak louder than words.” Nonverbal behavior is indeed largely automatic, meaning that very few people can consistently regulate subtle micromovements and gestures such as sidelong glances or genuine smiles. In this sense, nonverbal data are uniquely telling and definitely crucial for advertising, politics, and in general for identifying a person and their patterns.

Nonverbal behaviors are fundamental to creating immersive experiences in the Metaverse because they only work if the system measures body and eye movements to make the environment respond accordingly. For example, in VR, people turn their physical head around to make eye contact with other virtual reality users, use their legs to walk in the physical room to get across a virtual room, and move their physical arms to grasp virtual objects. These tracking data can be recorded and stored for later examination. In 2018, commercial systems typically tracked body movements 90 times per second to display the scene appropriately, and high-end systems recorded 18 types of movements across the head and hands. Consequently, spending 20 minutes in a VR simulation leaves just under two million unique recordings of body language (Bailneson, 2018). The largest part of the data collected can be included in three categories: hand tracking, eye tracking, and gait tracking, with the last two being the most interesting in terms of inferred data.

Eye-Tracking

Eye-tracking enables sophisticated data collection to tap into non-conscious processes governed by our biases and preferences. A combination of gaze position, pupil dilation, eye blinking rate, vertical and horizontal eye movements, and various other characteristics make eye tracking an attractive tool for qualitative and quantitative research.

Eye-tracking allows for the quantification of various key metrics that can be utilized to gain new insights and discoveries. Some of the common metrics that can be derived from eye-tracking include fixations and gaze points, heatmaps, Areas of Interest (AOI), Time to First Fixation, Time spent (Dwell time), Ratio, Fixation sequences, Revisits, First Fixation Duration, Average Fixation Duration. (*Understanding Human Behavior - A Physiological Approach - IMotions*, 2016)

Embedded eye-tracking sensors within these XR devices hold the keys to understanding the complex cognitive processes used during combat scenarios, allowing the simulation training to include cognitive performance metrics to better understand and improve the physical and mental responses of warfighters. Visual Recognition and Identification can be used to understand cognitive performance gains and the overall effectiveness of warfighters' training. Analyzing and evaluating the cognitive load and functions as well as characterizing deterioration over time provides insight into understanding and improving proper decision-making.

Scanning Pattern Proficiency and inferences gained from eye-tracking technologies provides insight into the cognitive effects of long-duration combat missions, executing multiple operations in one day, and recovering from tactical tasks. Military leadership can better assess the risk of accidents and fatalities under these conditions using the technologies and available data. Still, there needs to be a baseline understanding of what happened to all the data collected during these missions and exercises and where the data collection or sharing actually puts humans at risk. Massive, unchecked use of such technologies involving the use of BID increases the dangers to privacy, civil liberties, and free agency. This data in the hands of an adversary represents a significant national security concern.

Gait Tracking

Gait analysis is “the systematic study of animal locomotion, more specifically the study of human motion, using the eye and brain of observers, augmented by instrumentation for measuring body movement, body mechanics, and the activity of the muscles” (“Gait Analysis,” 2021). There are numerous uses for gait analysis, both good and bad and now many XR devices inherently collect data for gait analysis.

Traditionally conducted using external cameras, gait tracking has been very useful in diagnosing illnesses and issues that affect people's ability to walk. Having this data available, especially data that can show a change in gait of soldiers wearing combat loads in excess of existing limits, and without the use of an elaborate camera setup would help military planners better understand the consequences of adding additional weight to the already overloaded ground Soldier or Marine. Another way the data is valuable is in trend analysis. When collected over time, the data can show changes in an individual's health and can aid in proving service-connected disabilities for ground troops and speed up the Veterans Administration (VA) disability rating process. While camera systems are more reliable, instrumented floors and inertial measurement units (IMUs) consisting of accelerometers, gyroscopes, and magnetometers, are contributing data for analysis, and are becoming increasingly more reliable as gait analysis tools. In the case of XR devices, IMUs are mostly standard equipment used to help the system know where a user is looking, how fast they are moving, etc.

Gait data can also be very useful in accurately identifying people. Termed gait biometrics, studies have shown that the level of accuracy of identifying people is between 80-100% (Connor & Ross, 2018). The same study also pointed out that gait is hard, if not impossible to mimic, making it a potentially viable biometric authentication method. While an individual's gait can be temporarily faked (i.e., adding a fake limp), another person trying to accurately imitate another's genuine gait is near, if not completely, impossible. Because of this, gait may make a great biometric authentication method, especially for wearables and XR devices.

Intersections of XR and Other Emerging Technologies

XRSI has been observing the technology space, specifically, the intersections with other emerging technologies such as 5G/6G, BCI, AI, haptics, and multi-sensory. While new intersections continue to emerge and bring novel privacy and safety risks along with them, AI and BCI are the most concerning due to the excessive but necessary data collection.

Artificial Intelligence (AI)

Artificial intelligence is the study/domain of problem-solving, pattern recognition and developing systems armed with the intellectual characteristics of humans, such as the ability to reason, discover meaning, generalize, or learn from past experience. Currently, the AI-based systems identify objects but do not fully understand the context of inter-object relationships, whether it is hierarchical or flat, or interdependence of objects and their relationship with one another influencing events in real-time. The convergence of AI and XR will enable the next stage of assigning behaviors and understanding events by linking real-world with virtual objects within the Metaverse. This intersection of two evolving domains not only combines the associated risks but also exposes novel risks.

Metaverse stakeholders must prepare in advance to handle the evolving risks brought on by the convergence of new technologies such as AI and XR. While these technology convergences can be used to mitigate some of the emerging risks, they also introduce threats of their own. Liu et al. stated that “machine learning models, especially deep neural networks, have been recently found to be vulnerable to carefully-crafted input called adversarial samples.” (Liu et al., 2020) An example of an adversarial sample would be someone posting a picture on a stop sign that changes it to a different sign with the intent of crashing a self-driving car. This same kind of problem must be taken into consideration during the building process of any new XR system and experience. Likewise, just as humans have both conscious and unconscious biases, many of these artificial intelligence systems learn the same biases that humans have regarding gender, race, and attitudes toward people's differences. For example, gendered language can be introduced into these systems because machine learning models are only a reflection of the world in which we live (Hoyle et al., 2019). Recent innovations have shown that these biases can be mitigated proactively by conducting new research specifically to address these concerns.

5G/6G

To achieve the faster rates and lower latency performance gains, an efficient network will be required. The network must dynamically allocate resources, change traffic flow, and process signals in an interference-rich environment. 6G promises to offer persistent and immersive connectivity to humans, machines, and other devices [14] As with 5G, the focus of information transfer will continue to shift from humans to devices as the number of devices on the network explodes. As a data point, the current number of IoT devices is on the order of the human population worldwide [15]. Yet, by 2030 it is expected that the number of connected devices will be roughly 15 times the number of humans on the planet [16]. Increasingly, machines will need to be connected by means of wireless communications. Examples of connected machines include vehicles, robots, drones, home appliances, displays, smart sensors installed in various infrastructures, construction machinery, and factory equipment [17].

Brain-Computer Interface (BCI)

Sometimes called neural-control interface (NCI), mind-machine interface (MMI), direct neural interface (DNI), or brain-machine interface (BMI), a Brain-Computer Interface is a direct communication pathway between an enhanced or wired brain and an external device[19]. A BCI allows for bidirectional information flow opening up the possibility for “writing back” to the brain. BCIs are often used for researching, mapping, assisting, augmenting, or repairing human cognitive or sensory-motor functions.

BCI translates a user's brain activity directly into a computer control signal to activate a computer or other external devices. The brain signals are usually measured using electroencephalography (EEG) and processed by neural interfaces. A BCI in XR is ideally situated to supplement and improve conventional XR modalities, by both expanding the actions possible with XR, and providing a more flexible and immersive user experience overall. (Scherer et al., 2010)

BCI technology provides a solution to maximize the Metaverse's potential, affording users real-time mental selection via dry electroencephalography (EEG). The risks and benefits of these systems depend mainly on their level of invasiveness. This technology is by nature intrusive, and privacy and security risks must be assessed and mitigated as they emerge.

Unintended Consequences and Human Risk

As already discussed, the unprecedented quantity and variety of individuals' data collected and processed by XR devices is redefining the concept of anonymity. The amount of first-hand and inferred information that can be extracted from data can create unique and unmistakable patterns that permit identifying an individual no matter the anonymization of their data. Furthermore, data in aggregate can begin to show likes and dislikes, and people with access to the data can use that data – and the intimate (literally) “in your face” access to manipulate a person's behavior.

The design and color choices, button placement, and app behavior all lend themselves to encouraging people to act in a specific way. For instance, Disney's subsidiary Pixar Animation Studios uses specific colors at specific times to elicit emotions in movie scenes (Rogers, 2021). App developers strategically place (or hide) buttons and menu items to steer you down a specific path. Taken a step further, it is feasible for someone to combine these techniques above with personal information about an individual, injecting that into a visual environment for nefarious reasons. This injection of personal information can be timed and placed correctly to elicit emotions, foster trustworthiness, and ultimately sway decision a major decision (i.e., a Presidential election) or to create AI-based Agents capable of extracting sensitive information out of unwitting Service-members (Scharlat, 2007).

Traditionally, content and design considerations in digital media are left squarely in the domain of the creatives and developers, and little or no oversight is required in terms of safety and privacy when it comes to seemingly benign things like the background color of your webpage or which font you use. However, in the immersive domain, design considerations matter. If I suddenly remove the virtual floor from under a person and they start to fall, or don't, or the orientation of familiar objects like the sky are below and the ground is above you, the resulting disorientation can cause a physical reaction or injury.

The XRSI Framework

The XRSI Privacy and Safety Framework, also referred to as “The XRSI Framework” is a free, globally accessible baseline rulebook curated by the XR Safety Initiative, designed to provide the technical foundations for the regulatory guidance and build trust in the XR ecosystems aka the Metaverse.

The XRSI Framework is a community built, agile, iterative, and functionality-driven approach to safety and privacy in XR. The development process is managed by a steering committee composed of multidisciplinary professionals, from cybersecurity to medical experts, from developers to child safety advocates and more. The XRSI Framework creates a baseline set of standards, guidelines, and best practices that are regulation-agnostic. It addresses privacy and safety requirements drawn from the GDPR, National Institute of Standards and Technology (NIST) guidance, FERPA, COPPA, and a few other evolving bodies of guidance. The XRSI Framework is a living document designed to quickly adapt to new requirements, regulations, and approaches as soon as they come into effect. The framework is not a law or standard; it is a free tool that is continuously evolving.

Overview of The XRSI Framework

The XRSI Privacy and Safety Framework has four focus areas, set of function, and granular controls underneath. Each component reinforces how organizations and institutions achieve privacy and safety goals through aligning strategy, roles, and responsibilities.

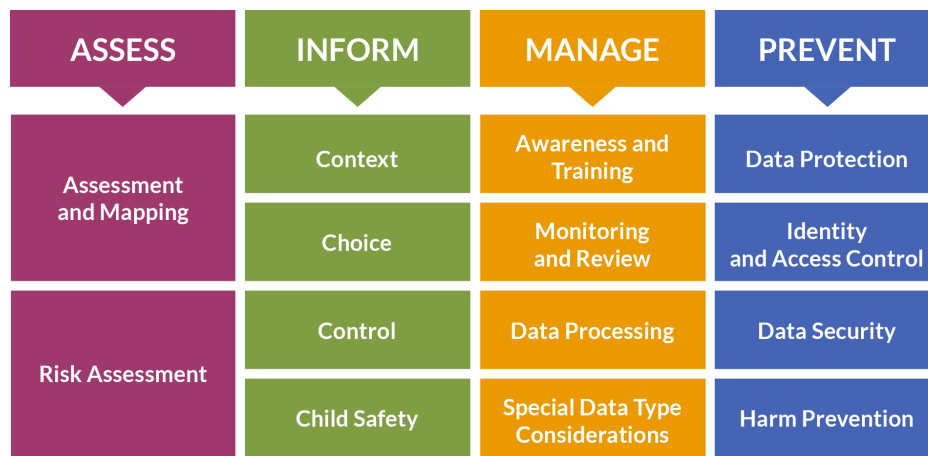


Figure 3 Key components and focus areas of the XRSI Framework

Focus Areas - The Focus Areas provide a foundation for outlining the scope of work that enable institutions to incorporate human-centric privacy and safety by design and default into their academic practices and development for the Metaverse.

Functions - Functions are the subcategories for outlining groups of privacy and safety-focused activities tied to the focus areas and corresponding granular controls.

Controls - Controls are the activities that are carried out to achieve specific privacy and safety outcomes about operations. They provide a set of results and help support the achievement of the intended outcome in each of the focus areas.

The XRSI Framework Focus Areas

The XRSI Framework is a voluntary tool for managing privacy and safety risks in immersive technologies and the Metaverse. It is intended to serve organizations and institutions of all sizes. The XRSI Framework is strategically designed to be compatible with existing domestic and international legal and regulatory regimes and usable by any type of organization to enable widespread adoption. It explicitly considers key regulations such as GDPR, CCPA, COPPA, FERPA, and a few others, as previously mentioned.

The XRSI Framework's purpose is to help organizations and institutions utilizing XR technologies, manage privacy and safety risks.

When used as a risk management tool, the XRSI Framework can help an institution build trust, achieve transparency, and create accountability during research, development, and innovation while minimizing unintended consequences for individuals or collectives. It can also help organizations and institutions exploring the Metaverse, to assess the scope of their impact on individual and collective rights, facilitate those rights, and comply with various international and state regulations.

The XRSI Framework is intended to be used not just as a compliance tool, but also to establish the baseline measure to optimize trust and safety efforts in minimizing risks within the XR Data Processing Ecosystems such as the Metaverse.

Assess

This area is focused on developing an understanding to manage privacy and safety risks for individuals and organizations arising from data processing and collection. Assessment of risks is fundamental to building responsible and human-centric Metaverse. Develop a comprehensive understanding of the organization's risks associated with data collection, processing, analysis, and its impact on the individuals. This enables an organization to understand the business environment in which it operates and prioritize risk mitigation accordingly.

Mapping and Analysis - Data processing systems, products, and/or services are mapped and analyzed for potential privacy and safety risks.

Risk Assessment - Privacy and safety risks related to the organizations are assessed to determine the impact on their operations, mission, and functions considering other risk factors, including human, societal, informational, financial, and legal risks.

PR – Prevent

The strong cybersecurity background within the XR Safety Initiative drives the Framework development with a clear understanding: preventing threats is always better than responding to them. More specifically, when this principle is plugged into the immersive domain, the prevention activities must be extended to a wide set of functions, such as:

- Data Protection
- Identity Management, Authentication, and Access Control
- Data Security
- Online Harm Prevention
- Child Safety.

Given the context, the nature, and the purposes of the processing of individuals' data (and the amount of personal data processed), XR providers should minimize any potential data/information exposure. Some XR providers do not ensure the adoption of certain data security measures, such as encryption or pseudonymization (standard practice in more traditional digital communication means such as instant messaging apps). Further, certain XR systems also rely on third-party services or apps that do not appear to implement suitable security standards. It is essential for the Metaverse stakeholders to implement adequate policies and security measures (e.g., physical security of data, physical security of facilities/personnel, network security, system hardening, password security, endpoint protection, patch management, remote access, etc.) to satisfy legal requirements. Protection of children in the Metaverse is a highly sensitive aspect: when it comes to minors, prevention is the only possible approach, because managing already performed threats could not be possible, and the harm could already be put into effect on a particularly vulnerable community.

MN – Manage

The Metaverse is evolving fast, and some of the products, devices, applications, and experiences are not XR-native. At the same time, new data processing tools and technologies are constantly released. This means that not every possible risk can be prevented, and the preventive approach must be integrated with a clear understanding of how to manage existing threats.

The MN area focuses on organizational-level activities such as establishing organizational values and policies, addressing safety, privacy, legal and regulatory requirements, managing organizational risk tolerance to allow an organization to focus and prioritize its efforts, consistent with its overall risk management strategy and business needs.

More specifically, the area has been divided into four critical functions, allowing organizations to understand and manage risks in the different phases of the product and data lifecycle:

- Awareness and Training: The organization's workforce and third parties engaged in data processing are provided privacy awareness education and are trained to perform their privacy-related duties and responsibilities consistent with related policies, processes, procedures, and agreements and organizational privacy values.
- Monitoring and Review: The policies, processes, and procedures for ongoing review of existing controls and risks to develop an effective plan of action.
- Data Disclosures (Breach Notification): Data breach notification requirements were designed to empower consumers and shame companies into improving their data security practices. The precise contours of when/where/how individuals must be informed varies widely based on jurisdiction.
- Data Processing Ecosystem Risk Management: The organization's priorities, constraints, risk tolerance, and assumptions are established and implemented to support risk management decisions within the data processing ecosystem.
- Special Data Type Consideration: The organization has established and implemented the processes to identify, assess, and manage privacy risks related to special data types. The organization's priorities,

constraints, risk tolerance, and assumptions are established and used to support risk decisions associated with sensitive data that may put humans at risk.

Inform

Articles 13 and 14 of the European Union's Law 2016/679, GDPR, states that any processing of personal data should be lawful, fair, and transparent. It should be clear and transparent to individuals that personal data concerning them are collected, used, or otherwise processed, and to what extent the personal data are, or will be, processed. The right to be informed, under Articles 13 and 14 GDPR, is a key part of any organization's obligations to be transparent.

The principle of transparency requires that any information or communication relating to the processing of personal data is easily accessible and easy to understand, and that clear and plain language be used. Any information addressed to the public or to the data subject must be concise, easily accessible, and easy to understand, and that clear and plain language and, additionally, where appropriate, visualization be used.

A New Definition of Personal Data

The Metaverse expands the definition of personal information that must be protected, including BID. The people whose information is tracked, collected, used, and shared, must be granted the right to be informed on the full data pipeline.

Given the potential immersion via the Metaverse and the breadth of sensitive information available to XR hardware, informed consent is critical. This concept becomes especially important when the data tracking, collection, usage, and sharing, involves vulnerable individuals or categories, such as minorities, persons with disabilities, or children. The right to informed consent includes ensuring age-appropriate design and awareness for parents to increase child safety.

The first crucial element an individual must be provided is context, meaning that any product or experience must clearly communicate what kind of data is managed, how the whole data funnel is shaped, and how long the data will be in every phase of this process.

In the XRSI Framework approach, the context is the minimum baseline for offering individuals a choice on what to share, how to share it, and with whom. This is not only a matter of development and programming (for instance, implementing a clear User Interface to provide opt-in and opt-out choices), but a cultural switch to a privacy-first and human-centric privacy-by-design approach, respecting privacy as a fundamental individuals' right.

Where to Find More Information

For more information about the Framework or how you can get involved, visit XRSI at www.xrsi.org or contact the authors directly via email.

Conclusion

The Metaverse is a new and emerging phenomenon and no regulations currently exist to mitigate safety and privacy risks in the Metaverse. This creates a massive gap in terms of our ability to trust in the technologies used to enable virtual experiences. Various ways of data collection such as eye and GAIT tracking combined with various technological intersections such as AI and BCI, multiply and significantly increase risks to humans, and society in general. While we await regulations and various technologies converge, the use of XR within critical sectors military and healthcare warrant us to adopt a self-governance model to address expected and unintended consequences.

With big technology companies moving fast to build and own the Metaverse, the XRSI Framework acts as the tool for assessing risks, informing individuals around the associated risks, managing the risks, and most importantly preventing harm with proactive technical, administrative and physical safety and privacy controls. Adopting the XRSI framework to establish baseline privacy and safety of XR ecosystems and the Metaverse will allow us to continue down the path of responsible and ethical innovation and even safeguard national security.

References

- [1] Bailenson JN. Experience on Demand: What Virtual Reality Is, How It Works, and What It Can Do. New York, NY: W. W. Norton; 2018.
- [2] Connor, P., & Ross, A. (2018). Biometric recognition by gait: A survey of modalities and features. *Computer Vision and Image Understanding*, 167, 1–27. <https://doi.org/10.1016/j.cviu.2018.01.007>
- [3] *Dark Patterns*. (n.d.). Retrieved June 13, 2021, from <https://www.darkpatterns.org/>
- [4] Gait analysis. (2021). In *Wikipedia*. https://en.wikipedia.org/wiki/Gait_analysis
- [5] Hornigold, T. (2019, March 7). OpenAI's Eerily Realistic New Text Generator Writes Like a Human. *Singularity Hub*.
<https://singularityhub.com/2019/03/07/openais-eerily-realistic-new-text-generator-writes-like-a-human/>
- [6] Hoyle, A. M., Wolf-Sonkin, L., Wallach, H., Augenstein, I., & Cotterell, R. (2019). Unsupervised Discovery of Gendered Language through Latent-Variable Modeling. *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, 1706–1716. <https://doi.org/10.18653/v1/P19-1167>
- [7] Kang, D. (2018, November 6). Chinese “gait recognition” tech IDs people by how they walk. <https://apnews.com/article/china-technology-beijing-business-international-news-bf75dd1c26c947b7826d270a16e2658a>
- [8] Reinhold Scherer, Mike Chung, Johnathan Lyon, Willy Cheung, and Rajesh P. N. Rao. 2010. Interaction with Virtual and Augmented Reality Environments using Non-Invasive Brain-Computer Interfacing. 2010 1st International Conference on Applied Bionics and Biomechanics (ICABB-2010), Venice, Italy, October 14-16, 2010 (2010).
- [9] *Research & Standards – The CyberXR Coalition 2021*. (n.d.). Retrieved June 13, 2021, from <https://cyberxr.org/research-standards/>
- [10] Rogers, A. (2021, April 29). *How Pixar Uses Hyper-Colors to Hack Your Brain | WIRED*. <https://www.wired.com/story/how-pixar-uses-hyper-colors-to-hack-your-brain/>
- [11] Scharlat, J. (2007). *Intelligent Virtual Environment Agents (IVEAs): Conducting Information Operations in Virtual Environments*. Naval Postgraduate School.
- Understanding Human Behavior—A Physiological Approach—IMotions*. (2016, March 28). Imotions Publish. <https://imotions.com/blog/understanding-human-behavior/>
- [12] *XRSI Privacy & Safety Framework* XRSI.org
- [13] *The Growing Value of XR in Healthcare in the UK* UK Research and Innovation NHS, Health Education England
- [14] F. Tariq, et al., *A Speculative Study on 6G*, *arXiv*, 2019.
- [15] Y. Kim, *Making 5G Wireless Technology a Reality and Initiatives Towards 6G*, 1st 6G Wireless Summit, <https://www.youtube.com/watch?v=2HKKE02SMnk>, 2019.
- [16] S. Choi, *6G Symposium Panel*, Samsung Research, 2020.
- [17] Samsung Research, *The Next Hyper-Connected Experience for All*, Samsung Research, <https://research.samsung.com>, 2020.
- [18] P. Mogensen, A. Ghosh, A. Maeder, M. Uusitalo and S. Redana, *5G Evolution and Beyond '6G?'*, 6G Flagship, http://www.6gsummit.com/2019/wpcontent/uploads/2019/04/Day3_Session5_Mogensen_Nokia-Bell-Labs.pdf, 2019.
- [19] *Research & Standards – The CyberXR Coalition 2021*
- [20] *Snow Crash* Neal Stephenson 1992
- [21] *Ready Player One* Ernest Cline 2011
- [22] *General Data Protection regulation (EU) 2016/679*,
<https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32016R0679#d1e2244-1-1>